Intégration de l'Information Stylistique dans le Dictionnaire Bilingue Spécialisé

Patrick Leroyer

The Aarhus School of Business
French Department and Center for Lexicography
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark
pl@asb.dk

Abstract

Integration of stylistic information in bilingual technical dictionaries has long been a major desideratum, as translators also need stylistic information about how professional discourses are conventionally realised. The lexicographic solution described in this paper is dedicated to satisfy stylistic information needs, and is compatible with the traditional terminological approach. It is realized by the lexicographic category of sentence examples as a sui generis category. First, the theoretical basis of the category will be outlined. The lexicographic encoding of the data will then be exemplified in the context of structural and functional dictionary integration. Finally, a number of new lexicographic rules will be stated for compilation of corpus, selection criteria and validation, lemmatization, treatment of collocation, article configuration, and distributional integration in and outside the macrostructure.

1 Introduction

Il est largement reconnu que la plus grande difficulté de la traduction technique réside dans la connaissance des termes utilisés par le spécialiste du domaine. C'est ce que Lerat [1995] appelle un préreguis incontournable, tout en notant que les langues spécialisées relèvent d'une culture de l'écrit. Les dictionnaires spécialisés dont l'objectif central est de fournir une assistance optimale à la traduction technique devraient ainsi être conçus comme des dictionnaires terminologiques, cf. [Bergenholtz 1994] et [Pedersen 1996], fournissant une couverture systématique du domaine traité, tout en utilisant des solutions adaptées aux compétences des utilisateur traducteurs, cf. [Tarp 1995]. Or, les traducteurs réclament aussi de leurs dictionnaires des informations illustrant la mise en discours des termes. Plus prosarquement, ils réclament des collocations, mais aussi des exemples phrastiques. Humbley [2000] constate que les solutions phraséologiques proposées jusque là n'ont pas vraiment su relever ce défi. Il est permis dès lors de penser que la terminologie, même en travaillant au niveau textuel, ne soit pas vraiment à même de fournir ce type d'information dans le dictionnaire, cf. en particulier [Bourigault/Slodzian 1999], ou bien que les structures du dictionnaire ne soient pas aptes à la supporter. La solution lexicographique décrite dans cet article est dédiée à l'intégration de l'information stylistique dans le dictionnaire, et s'appuie sur la théorie des genres textuels. Elle consiste à fournir deux types d'informations: des informations implicites dans les articles eux-mêmes, au moyen d'exemples phrastiques systématiquement sélectionnés, et des informations explicites dans un précis stylistique relié aux articles, et recyclant ces mêmes exemples. Un système de renvois multirelationnels complète la solution, et la rend particulièrement bien adaptée à la facture des dictionnaires électroniques.

2 La théorie du genre comme outil lexicographique

A quelques exceptions près, cf. [Harras 1986], [Nielsen 2000], et [Schneider 1998], les potentiels lexicographiques de la théorie du genre ont été sous-estimés par la lexicographie spécialisée. Pourtant, les dictionnaires basés sur le genre existent bien, même s'ils ne sont pas catégorisés comme tels dans les typologies classiques, cf. [Leroyer et al. 1999]. L'exemple le plus connu est sans doute celui du dictionnaire de correspondance commerciale, qui ne s'applique pas à délimiter et décrire le vocabulaire central de plusieurs domaines de spécialité, mais à délimiter et décrire les ressources lexicales et stylistiques du discours épistolaire, actualisé dans la lettre commerciale au travers des domaines. La théorie du genre, développée en particulier par [Bhatia 1996] et [Swales 1990], apporte une réponse satisfaisante aux propositions de [Mugdan 1989] et [Nielsen 2000]: le dictionnaire spécialisé devrait aussi contenir des informations sur la culture textuelle d'une profession. Le dictionnaire du domaine de spécialité devrait ainsi être complété par le dictionnaire de la communication privilégiée par le domaine. La différence ici est une simple différence de point de vue. Bessé [2000] le fait bien remarquer: le domaine n'a pas d'existence en soi, il est simplement délimité du point de vue de l'expert. C'est ainsi que les deux points de vue celui de l'expert, et celui du communicateur, le rédacteur technique par exemple – devraient aller de pair et se renforcer mutuellement en alliant représentation de la connaissance, et représentation de sa communication et de sa médiation. Le genre constitue ainsi une plateforme textuelle théorique débouchant sur deux types d'incorporation lexicographique.

2.1 Incorporation partielle

La figure 1 ci-dessous illustre le principe d'incorporation partielle du genre dans la conception du dictionnaire. Les exemples adressés aux lemmes et aux équivalents seront sélectionnés en fonction de la logique conceptuelle du système terminologique du domaine, mais ils seront aussi complétés par des exemples illustrant la mise en discours des termes, dans des genres fréquemment utilisés par la communication technologique du domaine. Un dictionnaire de l'énergie éolienne pourra ainsi inclure des exemples extraits de manuels de l'utilisateur, de brochures, de brevets, de communiqués de presse, etc. Ce type d'incorporation entraîne par conséquent une extension des critères de sélectivité lemmatique, et implique une réorientation pragmatique du dictionnaire.

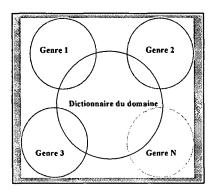


Figure 1: Incorporation partielle

2.2 Incorporation intégrale

La figure 2 ci-dessous illustre le principe d'incorporation intégrale. C'est désormais la logique du genre qui gouverne la conception du dictionnaire. Dans le cas du dictionnaire de correspondance commerciale par exemple, les exemples phrastiques adressés aux lemmes et aux équivalents illustreront les scénarios de la communication des affaires, qui sont basés sur une praxis conventionnalisée – l'offre, la demande, le règlement, etc. – et sur le recours à des actes de communication privilégiés. Les domaines concernés – banque, transport, comptabilité – permettront d'actualiser ces différents scénarios. L'incorporation intégrale du genre permet ainsi de cerner le vocabulaire essentiel de l'activité commerciale, d'en assurer la cohérence thématique, et surtout de représenter les stratégies rhétoriques privilégiées par le genre.

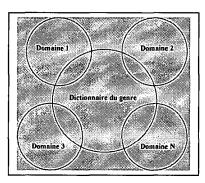


Figure 2: Incorporation intégrale

3 Sélection des exemples

 Les exemples stylistiques seront extraits d'un corpus traductique bilingue ou multilingue constitué pour représenter un domaine, ou un genre fréquemment exploité au travers des domaines. Leur incorporation ne sera pas exclusivement déterminée par la logique

- terminologique du domaine, mais par celle de la communication textuelle du domaine, c'est à dire du genre textuel.
- 2. L'accès lemmatique aux exemples, quelle que soit l'orientation du dictionnaire dictionnaire du domaine illustré par les genres, ou dictionnaire du genre illustré par les domaines sera déterminé par les résultats de l'analyse thématique des genres, et par la validation des candidats terminologiques mis en discours.
- 3. Les exemples stylistiques seront sélectionnés selon des critères de représentativité textuelle basés sur le découpage des discours techniques en scénarios thématiques prototypiques, et selon des critères collocationnels.
- 4. Les exemples stylistiques seront dotés de renvois à un précis stylistique et traductique externe à la macrostructure.
- 5. Pour les dictionnaires bâtis sur BD relationnelles, les exemples stylistiques seront indexés au moyen d'un triple paramétrage: domaine de spécialité, genre, et scénario thématique.
- 6. Les exemples stylistiques seront régis par un réseau d'adressage multirelationnel aux données distribuées au travers des structures dictionnairiques.

Le corpus traductique sera préférablement constitué de documents informatisés (format pdf par exemple). Les documents seront tirés de sites reconnus pour la qualité des traductions multilingues. Le corpus pourra également être basé sur des documents mis à la disposition du lexicographe, cf. [Humbley 2000]. Les exemples seront validés conformément à Schneider [1998], avec l'intervention de spécialistes de la traduction en vue de valider la qualité traductique.

4 Configuration d'un article prototypique

La figure trois ci-dessous illustre les principes appliqués à la structuration et la formalisation microstructurelle des exemples stylistiques dans un article prototypique, pour le couple français danois.



Figure 3: Exemple d'article prototypique

Les catégories d'indication et les solutions de marquage sont les suivantes:

- Indication du lemme
- Indication grammaticale (conjugaison) avec marqueur de renvoi * aux tables de verbes irréguliers
- Indication de la classe de mot (partie du discours)

- Indication mono- ou multi-équivalents listés en ordre alphabétique sans discrimination sémantique (principe de sous-détermination de la traduction au niveau de la substitution)
- Indication de l'exemple source avec adressage lemmatique interne
- Indication de l'exemple cible avec adressage multi-équivalentiel interne
- Indication de l'exemple source et cible avec adressage relationnel du type un à plusieurs des données stylistiques (réexploitation d'un même exemple au travers du précis)
- Marquage typographique du sens traductique: caractères romains pour l'exemple source, et caractères *italiques* pour l'exemple cible
- Marquage contrastif ≠ indiquant la divergence des choix traductiques
- Champs de renvois aux items du précis stylistique avec marqueur de renvoi ⇒

5 Méthode

5.1 Méthode intégrative

La solution lexicographique décrite dans cet article est une solution authentiquement intégrative, car bâtie sur l'instauration de rapports endogènes gérant les relations d'adressage interne entre composantes dictionnairiques fonctionnelles et structures d'accès externes. Traditionnellement, en méthode linguistique, la démarche syntaxique descriptive est basée sur l'analyse théorique des données empiriques illustrée par le regroupement systématique en "catégories d'exemples". De manière similaire, la représentation lexicographique des données syntaxiques fonctionnelles (desservant la fonction traduction) est fondée sur le recyclage de la méthode syntaxique. Mais à ceci près que les données syntaxiques sont simultanément dotées d'un adressage fonctionnel de type direct/indirect: les données centrales utiles à la traduction, c'est à dire les équivalents, et les données périphériques utiles à la connaissance de la traduction, c'est à dire le commentaire stylistique de type contrastif. Les données lexicographiques sont prises en charge par une double configuration relationnelle bidirectionnelle, du type un à plusieurs, et plusieurs à un, puisque chaque exemple est relié à plusieurs articles du précis stylistique, alors que les articles du précis sont reliés à plusieurs exemples.

5.2 Méthode descriptive

Conformément à Lerat [1995], la syntaxe privilégiée par les textes spécialisés est révélatrice de choix stylistiques qui sont conformes aux normes et aux conventions textuelles exploitées par les genres textuels au sein des discours professionnels. Plutôt que de parler de précis grammatical ou syntaxique intégré au dictionnaire, suivant les recommandations de Mugdan [1989], il vaut mieux parler de précis stylistique et traductique. En effet, ce qui est attendu de l'information syntaxique dans le dictionnaire — outre l'information fournie par les collocations — c'est l'aptitude à rendre compte lexicographiquement d'habitudes d'expression statistiquement dominantes dans tel type de texte et dans tel genre, c'est à dire d'un style, en adéquation avec la fonction dictionnairique centrale. Cette aptitude sera prise en charge par les exemples stylistiques, et par les commentaires qui leur sont adressés. Les données stylistiques seront de type implicite dans les articles, et desserviront la fonction directe du dictionnaire en apportant à l'utilisateur des informations illustrant l'usage des lemmes et de leurs équivalent dans des contextes actualisés. Par contre, les données stylistiques seront de type indirect dans le précis, puisqu'elles apporteront des informations

stylistiques contrastives dotées de commentaires. Dans les deux cas toutefois, les informations stylistiques fournies par le dictionnaire seront adressées à la fonction traduction du dictionnaire. Les analyses syntaxiques de type contrastif, qui sont sous-jacentes aux commentaires stylistiques, réutiliseront les méthodes valentielles développées par Herslund [1996], puisque l'approche valentielle se prête bien non seulement à la description des phrases, mais aussi à l'explication modulaire des différents niveaux de leurs constructions: sémantique, syntaxique, et pragmatique.

6 Construction du précis

L'accès externe aux données du précis sera assuré par la lemmatisation des phénomènes syntaxiques, qui sont identiques, pour chaque phénomène considéré, aux items intégrés dans les champs de renvoi des articles. L'accès interne sera assuré par quatre champs consultatifs: un champ définitoire, un champ explicatif traitant les exemples incorporés dans le précis, et deux champs de renvoi. Le premier renvoi est adressé aux entrées de la liste lemmatique: il permet de consulter les exemples syntaxiques illustrant le même phénomène syntaxique, mais non incorporés dans le précis. La solution de renvoi lemmatique à partir des items du précis permet d'alléger la rédaction du précis, et de limiter son volume. Elle permet aussi d'étendre l'application des données du précis au plus grand nombre d'exemples nécessaires pour couvrir les réalisations prototypiques du phénomène syntaxique considéré au travers du dictionnaire. Le second renvoi est adressé à des sources documentaires externes au dictionnaire, permettant à l'utilisateur d'approfondir les informations fournies par le dictionnaire.

Accès interne au sein des articles du dictionnaire

L'accès interne au précis sera pris en charge par l'adressage individuel du ou des phénomènes syntaxiques considéré(s) à chacun des exemples textuels. Les exemples démontreront des facteurs heuristiques exploités dans la rédaction du précis stylistique et traductique. L'adressage sera matérialisé par un ou plusieurs marqueurs de renvoi explicite(s) aux entrées lemmatisées du précis.

Adressage relationnel des entrées lemmatiques du précis

Les entrées lemmatiques du précis seront toutes adressées aux exemples textuels syntaxiques correspondants et intégrés dans le texte des articles du dictionnaire, à l'exception toutefois des entrées lemmatiques constituant des articles de renvoi interne. L'adressage relationnel des entrées du précis sera du type un à plusieurs, dans la mesure ou un article du précis renvoie à plusieurs exemples textuels distribués au travers de la microstructure du dictionnaire.

Aide à la rédaction des définitions linguistiques

La rédaction des définitions linguistiques adressées aux termes linguistiques pourra se faire sur la base des définitions fournies par les dictionnaires linguistiques, ou par les dictionnaires monolingues de référence, incluant ce type de termes, et en fournissant la définition (cf. [Bergenholtz 1997] pour l'étude des dictionnaires linguistiques et lexicographiques dans les pays scandinaves).

Déclaration informative

Le précis stylistique et traductique sera présenté dans la déclaration informative du dictionnaire (généralement imprimée au dos de l'ouvrage). Pour la version informatisée, la déclaration devra être intégrée au menu principal, ou à la page d'accueil dans le cas d'un dictionnaire accessible sur le Web.

Fonction dictionnairique centrale du précis

La fonction dictionnairique centrale du précis sera de desservir la fonction traductique périphérique du dictionnaire de traduction technique. Il s'agit donc d'une fonction conforme à la typologie des fonctions tarpiennes liées à la connaissance, dans laquelle l'acte de consultation du précis est motivé par le besoin d'accéder à un corps de connaissance, et à la récupération de savoirs stylistiques et traductiques.

Lemmatisation

Les termes syntaxiques et stylistiques seront lemmatisés selon leur forme lexémique canonique, sans recours d'aucune sorte aux abréviations. Cette règle s'appliquera également à la lemmatisation des unités polylexicales figées.

Exemples d'entrées

cataphore, construction présentative, focalisation, métataxe

Lemmatisation des termes linguistiques difficiles

Les termes linguistiques susceptibles de poser des obstacles à la compréhension de l'utilisateur, mais qui seront néanmoins employés dans la rédaction des commentaires du précis, seront lemmatisés et intégrés dans la liste lemmatique du précis. Ces termes linguistiques lemmatisés seront suivis d'une brève définition fonctionnelle, et seront exemplifiés. Ils seront marqués en caractères italiques dans les textes du précis, pour indiquer la présence d'un renvoi interne relié à l'entrée de l'article concerné. En version électronique, ces renvois internes pourront être gérés par des hyperliens vers l'intérieur de la macrostructure du précis, qui accélèreront la vitesse de consultation du précis, et augmenteront la flexibilité de son utilisation. Certains termes linguistiques par contre constitueront un simple article de renvoi, comme les deux articles ci-dessous. Cette solution permettra de regrouper les données dans des macrocatégories syntaxiques (modalité par exemple), et de limiter ainsi le nombre d'articles du précis, sans pour autant en compromettre son degré de précision, ni affecter l'accès de l'utilisateur à l'ensemble des données: déontique ⇒ modalité, épistémique ⇒ modalité

Structure textuelle des articles du précis

Tous les articles du précis seront bâtis autour d'une distribution textuelle quadripartite des données lexicographiques, qui sera disposée dans l'ordre suivant:

- 1. Une brève définition fonctionnelle du phénomène syntaxique privilégié avec exemplification.
- 2. Un commentaire syntaxique et stylistique contrastif des exemples des articles du dictionnaire (principe de recyclage lexicographique).

- 3. Des renvois alphabétisés aux items de la liste lemmatique permettant de consulter les exemples qui ne font pas l'objet d'un commentaire, mais démontrent et illustrent les mêmes caractéristiques syntaxiques (principe de recyclage partiel du fonds global d'exemples textuels). Les renvois explicites seront marqués par l'indicateur "⇒".
- 4. Des renvois à des articles et des ouvrages syntaxiques traitant du phénomène considéré. Les renvois seront marqués par l'indicateur "⇒".

Modalités d'accès externe

L'accès externe aux données lexicographiques du précis sera pris en charge par une liste lemmatique conventionnelle, à alphabétisation droite.

Mode d'emploi

Les modalités d'utilisation détaillées du précis seront brièvement décrites dans l'introduction métalexicographique du dictionnaire. Le précis sera précédé d'une fiche déclarative des spécifications typographiques, et des caractéristiques des différents renvois et marqueurs de renvoi.

Rédaction du précis

Le précis sera rédigé dans un métalangage définitoire et explicatif clair et simplifié, évitant les longueurs, et qui mette en évidence l'importance traductique des phénomènes syntaxiques contrastifs sélectionnés. La qualité pédagogique de la rédaction constitue un critère central de fonctionnalité. Dans la version papier du dictionnaire, le volume du précis mesuré en nombre de pages ne devra pas excéder la proportion de 15% maximum du nombre de pages global du dictionnaire, afin de ne pas surcharger et déséquilibrer inutilement la mégastructure de l'ouvrage, et de ne pas contraindre les marges d'extension optimales de la composante centrale: la liste lemmatique. Dans la version informatisée par contre, le volume du précis pourra être largement augmenté, en particulier avec un nombre beaucoup plus important d'exemples commentés (recyclage à plus grande échelle), sans que cela risque d'alourdir l'architecture dynamique du dictionnaire ou d'affecter sa convivialité, ni de limiter pour autant les ressources de la liste lemmatique. La capacité de stockage des données sera déterminée en concertation avec le concepteur de la base de données.

Références bibliographiques

Le précis stylistique et traductique sera complété par une liste détaillée des références bibliographiques adressées aux renvois du type 4. En version informatisée, certains renvois pourront être reliés à des sites grammaticaux internet traitant le ou les phénomène(s) considérés.

7 Conclusion

Confrontés aux contraintes d'espace des dictionnaires papier, les lexicographes spécialisés ont délaissé les exemples phrastiques comme catégorie d'information, par crainte de gonfler le volume de leurs ouvrages. L'argument n'est pas tout à fait convaincant, car il faut bien se rendre compte, comme le signale Pruvost [2000], que l'espace limité est justement ce qui a permis à la lexicographie de développer la force et la spécificité de sa méthode. Toujours contraint de se limiter, le lexicographe se doit de sélectionner les données qui répondent le

mieux à la finalité de l'ouvrage envisagé. La solution décrite ci-dessus s'applique sans difficultés aux dictionnaires papier. Elle permet en effet de supprimer l'espace traditionnellement alloué aux collocations, en incorporant celles-ci dans les exemples stylistiques eux-mêmes. Pour les dictionnaires électroniques, la solution révèle toute sa force. Elle permet d'indexer l'ensemble des exemples stylistiques, et de les relier aux entrées du précis. La mise en place d'hyperliens permettrait en outre à l'utilisateur de consulter les informations distribuées au travers des structures du dictionnaire, et aussi de consulter directement des ressources externes venant compléter ces informations. Quel que soit son suppport, l'intégration des exemples stylistiques permet surtout de réorienter le dictionnaire bilingue spécialisé dans le sens des discours professionnels, tout en assurant la nécessaire compatibilité avec les méthodes terminographiques, et ouvre donc la voie à des dictionnaires de la communication technique.

8 Références bibliographiques

- [Bergenholtz 1994] Bergenholtz, H., 1994. Zehn Thesen zur Fachlexikographie. In: Burkhard Schaeder/Henning Bergenholtz (Hrsg.): Fachlexikographie. Fachwissen und seine Repräsentation in Wörterbüchern. Tübingen: Gunter Narr Verlag, 43-56.
- [Bergenholtz 1997] Bergenholtz, H., 1997. Linguistische und lexikographische Fachwörterbücher in Skandinavien. In: *Lexicographica* 13, International Journal for Lexicography. Tübingen: Max Niemeyer Verlag, 3-16.
- [Bergenholtz 2001].Bergenholtz, H., 2001. Lexikografie ist Selektion ist Selektion... In Korhonen, Jarmo (Hrsg.): Von der mono- zur bilingualen Lexikografie für das Deutsche. Frankfurt am Main: Peter Lang, 11-30.
- [Bessé 2000] Bessé, B. de, 2000. Le domaine. In: Le sens en terminologie. Lyon: Presses Universitaires de France, 182-197.
- [Bhatia 1993] Bhatia, V. K., 1993. Analysing Genre: Language Use in Professional Settings. London: Longman.
- [Bourigault/Slodzian 1999] Bourigault, D./Slodzian, M., 1999. Pour une terminologie textuelle. In: *Terminologies nouvelles*, no 19. Bruxelles: Agence de la francophonie et Communauté française de Belgique, 29-32.
- [Harras 1989] Harras, G., 1989. Zu einer Theorie des lexikographischen Beispiels. In: Wörterbücher. Dictionaries. Dictionaries. Ein internationales Handbuch zur Lexikographie. An International Encyclopedia of Lexicography. Encyclopedie internationale de lexicographie. Erster Teilband, hrsg. von Franz Josef Hausmann, Herbert Ernst Wiegand, Ladislav Zgusta. Berlin/New York: Walter de Gruyter, 607-714.
- [Humbley 2000] Humbley, J., 2000. Nouveaux dictionnaires, nouveaux rapports avec les utilisateurs. A paraître début 2002 dans *Meta*. Montréal: Les Presses de l'Université de Montréal.
- [Lerat 1995] Lerat, P., 1995. Les langues spécialisées, Paris: Presses Universitaires de France.
- [Leroyer et al. 1999] Leroyer, P., 1999. Dansk-fransk erhvervsordbog. Århus: Systime.
- [Mugdan 1989] Mugdan, J., 1989. Grundzüge der Konzeption einer Wörterbuchgrammatik. In: Hausmann, F.J./Reichmann, O./Wiegand, H. E./Zgusta, L. (Hrsg.): Wörterbücher, Dictionaries, Dictionnaires. An international Encyclopedia of Lexicography, First Volume. Berlin/New York: Walter de Gruyter, 732-749.
- [Mugdan 1998] Mugdan, J., 1989. Grammar in Dictionaries of Languages for Special Purposes (LSP). In: *Hermes*, Journal of Linguistics no. 3. Århus: The Aarhus School of Business, 125-142.

- [Nielsen 2000] Nielsen, S. 2000. Translation Strategies for Culture-Specific Textual Conventions in Bilingual Dictionaries. In: Lexicographica 16/2000, International Journal for Lexicography. Tübingen: Max Niemeyer Verlag, 152-168.
- [Pedersen 1996] Pedersen, J., 1996. Koncernspecifik brancheordbog med særligt henblik på teknisk engelsk. Prækonceptionnelle overvejelser og koncept. In: *Hermes*, Journal of Linguistics no. 16. Aarhus: The Aarhus School of Business, 267-275.
- [Pedersen 1998] Pedersen, J., 1998. Kollokationelle og syntaktiske angivelser i en teknisk oversættelses- og produktionsordbog for fremmedsprogsbrugere. In: *LexicoNordica* 5. Oslo: Nordisk forening for leksikografi, 53-65.
- [Pruvost 2000] Pruvost, J., 2000. Dictionnaires et nouvelles technologies. Paris: Presses Universitaires de France.
- [Schneider 1998] Schneider, F., 1998. Studien zur kontextuellen Fachlexicographie. Das deutschfranzösische Wörterbuch der Rechnungslegung. Lexicographica/Series major 83. Tübingen: Niemeyer.
- [Swales 1990] Swales, J., 1990. Genre Analysis English in Academic and Research Settings. Cambridge: Cambridge University Press.
- [Tarp 1995] Tarp, S., 1995. Wörterbuchfunktionen: Utopische und realistische Vorschläge für die bilinguale Lexicographie. In: Herbert Ernst Wiegand (Hrsg.): Studien zur zweisprachigen Lexikographie mit Deutsch II. Germanische Linguistik 127-128. Hildesheim/New-York: Olms, 17-61.